DOCUMENT RESUME

ED 077 948 TM 002 754

AUTHOR	Harman, Harry H.
TITLE	An Evaluation of the Multivariate Methodology of the Project.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RM-73-4
PUB DATE	Mar 73
NOTE	14p.; Paper presented at a symposium of annual meeting of the American Educational Research Association (New Orleans, Louisiana, February 26, 1973)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Cognitive Tests; *Concept Formation; Correlation; *Factor Analysis; Speeches; *Statistical Analysis; *Test Construction

ABSTRACT

Presented at a symposium on "The Structure of Concept Attainment Abilities Project: Final Report and Critique," this paper provides the methodological aspects of the project. The discussion centers around a "Guide to the Multivariate Methods," which is provided in the paper. The basic guide-posts are the types of analysis and the types of content. The latter include concept attainment by subject-matter field or combined, and cognitive abilities, or both. The three factor analytic techniques used to examine the data and obtain derived factor solutions. The next major phase of the analysis, the construction of 56 cognitive tests for the 1970 study, the factor analysis of these tests, and the reduced battery for use in the 1971 testing, is briefly noted. The most important part of the project--the study of relationships between the concept attainment measures and the cognitive abilities measures--is then described. The interbattery approach of Tucker, used to determine if there were factors common to the battery of cognitive abilities tests and the battery of concept attainment measures, is discussed. (DB)

RM-73-4

RESEARCH MEMORANDUM

AN EVALUATION OF THE MULTIVARIATE METHODOLOGY OF THE PROJECT

Harry H. Harman

This paper was presented at the annual meeting of the American Educational Research Association, Symposium on "The Structure of Concept Attainment Abilities Project: Final Report and Critique," New Orleans, February 26, 1973.

Educational Testing Service
Princeton, New Jersey
March 1973

An Evaluation of the Multivariate Methodology of the Project

Harry H. Harman

While my evaluation is limited strictly to the methodological aspects of the project, I do want to admit my acquaintance with and interest in the more general aspects of the project. Further, my knowledge and evaluation of the methodology is not restricted to the foregoing papers. Not only did I receive the 500 page draft copy of the monograph by the Harrises [7], but I also had the pleasure of being chairman and discussant two years ago of an AERA Symposium in which the Harrises' paper on the classification of cognitive abilities [6] gave an early report on that part of the project (Chapter IV).

Before I home in on the multivariate methodology let me give a very brief overview of my perception of the total project. In effect a type of systems analysis was made of four subject-matter fields, identifying the concepts that set these fields apart (Chapter II). Then measures (tests) were developed for the determination of the degree of attainment of these concepts, both in terms of "content" (i.e., the concept itself) and in terms of the "level of understanding" of the concept (designated "tasks"). In order to get some indication of the potential predictability of this type of achievement (i.e., concept attainment), the investigators explored several classification systems for cognitive abilities in building their own test battery. A variety of analyses were performed in the separate studies of the concept attainment measures (Chapter III) and of the cognitive abilities measures (Chapter IV), with the culminating analyses involved in studying the relationships between the two sets of variables (Chapter V).

Now, we turn our attention to an examination of the data summary and analysis. It would be the understatement of the day to say that this ambitious undertaking yielded a great wealth of results. The goals of the study were attained, in large measure, through the very effective planning and guidance of the analysis provided by Mr. Harris.

The first problem noted in his paper is not of the usual multivariate type. Basically, the question is how best to use the item data on the completely crossed design of concepts by tasks in constructing the tests of concept attainment. I checked with item-analysis specialists at ETS. When it became evident that they had not had previous experience with this type of data and did not have any immediate recommendations for its resolution, I could only agree with Harris on his approach--separate scores by rows and columns and proceed with traditional item analysis. But I also concur with him that this area merits further exploration.

The remainder of my discussion will be concerned with the more familiar multivariate problems of the project. I hadn't proceeded very far into this area before I saw the need for a map of this vast terrain. I share this with you in the form of a "Guide to the Multivariate Methods" in the Handout. The basic guide-posts are the types of analysis in the rows, and the types of content in the columns. The latter include concept attainment by subject-matter field or combined, and cognitive abilities, or both. A particular locale is found according to the analysis-and-content coordinates and is marked by a symbol that denotes concepts, tasks, or cognitive abilities, or their combinations, with the table numbers from the draft monograph where the data can be found. Also shown on the map are lines indicating the course of flow among the various forms of analysis.

The investigators rely heavily on factor analytic techniques in the examination of the voluminous data--matching my personal predilection. A whole range of problems is attacked with factor analysis, from exploratory efforts aimed at data reduction and determination of structure to the study of relationships between batteries of tests. The factor analysis work is sophisticated and pertinent to the objectives.

They always employ for the direct (or initial) factorization of a correlation matrix, a method due to Kaiser and Caffrey (Alpha); a method developed by Harris (referred to as $R\text{-}S^2$); and a maximum likelihood method of Jöreskog. The solutions yielded by all three of these methods have the property of being independent of the original scale or metric of the variables. That is why Harris selected these procedures, and it is a perfectly sound basis.

Both Alpha and Harris' method employ as a point of departure Rao's canonical factor analysis [10] in which the canonical correlations are determined between the observed variables and estimated factor scores. But they depart from Rao's statistical criterion for the number of common factors; Kaiser and Caffrey use the notion of psychometric generalizability while Harris uses the squared multiple correlation for estimating communality, in the sense of Guttman's best lower bound, for determining the number of common factors. Furthermore, in order to assure that the solutions are scale-free, in Harris' method the variables are rescaled in the metric of the unique parts while in Alpha they are rescaled in the metric of the common parts, and upon conclusion of the factoring the results are transformed back into the original metric of the variables.

Upon rotation of the initial solutions, Harris noted that the derived solutions based on Alpha factor analysis usually yielded only one factor while his method produced the largest number (with maximum likelihood producing an intermediate number). He also surmised that solutions with several factors probably were overdifferentiating because the derived oblique factors tended to correlate highly. The inclination of the project staff to follow a conservative course in interpreting a minimal number of factors agrees with my personal bias to err on the side of under-factoring--a willingness to have the factor analysis provide a "first approximation" model for the empirical data.

Before continuing with a discussion of derived solutions one might ask, why three separate methods? Several different factorizations are required for the application of the Harris' strategy for factor interpretation. That strategy [5] calls for derived solutions (both orthogonal and oblique) based on several initial factorizations, and from the results to accept as "the important substantive findings those factors that are robust with respect to method," i.e., factors that tend to include the same variables across method. These they call "comparable common factors" or CCF's. It should be clear that their concern is with potential idiosyncrasies of particular methods that may lead to unwarranted substantive conclusions. The strategy does not disclose the effect of chance errors in the data. It occurred to me that if they were to become concerned with sampling problems, they might employ a procedure recently applied to factor analysis by Pennell [9] that has been advanced by Tukey [13] as the "jack-knife," named for the boy scout's rough-and-ready general purpose instrument.

Returning to their use of derived factor solutions, we note from the Guide that they rely on three types: one orthogonal and two oblique. There is no question but that varimax is generally accepted as the preferred orthogonal solution. For oblique solutions, the issue is not so clear. The methods used were developed by Harris and Kaiser [4] in their 1964 paper, "Oblique factor analytic solutions by orthogonal transformations." Two of these procedures were used in the project, namely: (1) a method in which the reproduced correlation matrix can be represented by a set of independent (but correlated) clusters; and (2) a method in which the minor product moment of the factor pattern matrix is approximately proportional to the matrix of factor correlations. These two oblique methods are designated "Independent cluster" and "A'A proportional to L," respectively. These methods were compared with several other techniques for oblique transformations by Hakstian [1], who demonstrated in 1971 (and has provided further evidence at an AERA program this morning [2]) that they produce solutions best exemplifying simple structure.

The "independent cluster" factor solution can be expected to fit only simple data while the "A'A proportional to L" transformation can fit more complex data. Hence, it is not surprising that they found they had to reject the hypothesis of independent clusters after obtaining such solutions. The more complex oblique solutions were obtained from each of the three initial factorizations and interpreted by means of the comparable common factors strategy, as indicated by the flow lines into the CCF boxes in the Guide.

In studying the concepts and tasks for the four subject-matter fields, <u>separately</u>, the numbers of factors in the several solutions varied considerably, making the CCF strategy inappropriate. On the other hand, when they performed factor analyses of the concept measures and of the task measures for the four fields <u>combined</u> they found their interpretation strategy to be very effective. As noted by Mrs. Harris, this led to the overall conclusion of a comparable common factor representing each of the subject fields (with some slight overlap) for the concepts and for the tasks, as shown at the bottom of Table 1 in her Handout. Apparently the analysis required data revealing the contrasts among the four subject-matter fields to identify factors with fields and distinguish common from unique variance clearly.

For the analysis of concept attainment, there remains the important question of possible concept-task interaction in the crossed design of their study. This was the second problem to which Mr. Harris referred in his paper. It seems quite reasonable that they turned to Tucker's three-mode factor analysis [12] for answers to this problem. This procedure was applied to the 1970 data for each of the four subject-matter fields, and aside from making some necessary compromises, they found that there were no important concept-task interactions. Having reached this conclusion for the 1970 data, they did not repeat the analysis for the 1971 data. While they appeared pleased with the potentialities of three-mode factor analysis they also found its current computer programs limited in several respects.

The next major phase of analysis in the project deals with the dimensions of a battery of cognitive abilities tests, as shown in the two right-most columns of the Guide. First they reviewed the Guilford, Guttman, and Thurstone schemata for classifying cognitive abilities. This led them to construct 56 tests for the 1970 study. Upon factor analyzing these, and using their comparable common factors strategy, which worked quite well in this case, they determined the factors that they wanted preserved in a reduced battery for wider use in the 1971 testing (summarized in Table 2 of Mrs. Harris' Handout).

As Harris noted in his paper, the problem of selecting the subset of tests, indicated by the feedback loop at the bottom of the last column in the Guide, is not a trivial one. Their approach relied on the coefficients of the oblique factors (i.e., pattern matrix) in deciding whether to select a particular variable as a measure of a factor. But these regression coefficients "go the other way." They might have obtained the regression estimate of an oblique factor on the set of variables and thereby have the regression coefficients "go the right way" to facilitate the selection of important tests for a factor.

All the preceding analyses, in a sense, were preparatory for the most important part of the project--the study of relationships between the concept attainment measures and the cognitive abilities measures. Four essentially different attacks on this problem were considered, namely:

(1) Conventional factor analysis of the two batteries, simultaneously, without making a distinction between them;

(2) On some basis, designating one battery as "dependent" and projecting its variables into the common factor space of the other or "independent" battery;

(3) Canonical correlation and canonical variates approach;

(4) Interbattery factor analysis.

Although they performed factor analyses of the concept attainment measures and cognitive abilities measures treated as a single battery, they found the results to be less valuable than those obtained from the interbattery approach of Tucker [11].

While the interbattery procedure was designed originally to determine the stability of factors in two batteries of tests (assumed to depend on the same factors), it was used in this project to determine if there were factors common to the battery of cognitive abilities tests and the battery of concept attainment measures (without prior design for these two types of measures to depend on the same factors). The specific modifications they made was in using a different statistical test for the number of factors and by employing the interbattery factor matrices for <u>both</u> test batteries in getting an orthogonal and the two types of oblique derived solutions.

The number of interbattery factors was determined, in part, from the number of significant canonical correlations. Another use of canonical variate analysis was in getting squared multiple correlations of the concept attainment measures as estimated from the 31 cognitive abilities tests. In addition, they computed squared multiple correlations (i.e., communalities) of the concept attainment measures as estimated from the interbattery factors.

As noted above, four different attitudes about the two batteries were considered for the study of relationships between them. In ruling out one of them, they asserted that an approach involving canonical correlation and canonical variates implies a component-type solution. But that need not be the case. In a method developed by Bary Wingersky [14] at ETS, such a procedure is employed in getting a conventional factor solution. The method leads to complete and consistent solutions by satisfying the following two conditions (which are not exclusively satisfied by components solutions):

(1) Orthogonal factor scores are derived from the observed data and the factor pattern by a least-squares fit of the original data to its reconstruction from the factor model;

(2) The regression of the original data on these factor scores reproduces the factor pattern.

It should be noted, of course, that while the factor scores are implicit in the theoretical development, they are not actually computed for the cases in the sample. To avoid components solutions, each set of factor scores is the canonical variate associated with a canonical correlation that explains the most variance over all the data. Thus canonical correlations guide decisions concerning the number of factors to select, but in terms of reliability (as in Alpha factor analysis) rather than in a statistical sampling sense.

In the normal application of this method, a partition of the data is selected for the calculation of canonical correlations that will produce the largest estimate of reliability. I thought of adapting this method to the study of the relationships between the two batteries of the project by *a priori* assignment of them to the two partitions. The specific example used was data for the 1971 boys; the battery of 31 cognitive abilities tests was put in one partition and the battery of 48 combined subject-matter tasks was put in the other partition. Two and three factor solutions were obtained for comparison with the Harrises' results, although the new interbattery method pointed to possibly four reliable factors for the two batteries. Derived orthogonal (varimax) and oblique (A'A proportional to L) solutions were also determined for the two factor case since that is the number of interbattery factors exhibited in Table 8 of Mrs. Harris' Handout. No attempt will be made here to compare the results of the two approaches to interbattery factors. Suffice it to say that the squared multiple correlations for estimating the 48 task measures by the new method exceeded those of the original method by an average of only .03.

In conclusion, I want to thank the Wisconsin Research and Development Center for Cognitive Learning for inviting me to comment upon the methodology of this project that has, for all its complexity, been so well organized and carefully executed. Beyond its important substantive contribution to our understanding of how cognitive abilities combine in subject-matter concept attainment, the design of this project, its sensitive and sensible adaptation of methodology to objective, is a paradigm for what can be achieved with multivariate research techniques.

References

1. Hakstian, A. R. A comparative evaluation of several prominent methods of oblique factor transformation. Psychometrika, 1971, 36(2), 175-193.

2. Hakstian, A. R. Recent developments on the oblique transformation problem. Paper presented at the meeting of the American Educational Research Association, New Orleans, February 1973.

3. Harris, C. W. Some Rao-Guttman relationships. Psychometrika, 1962, 27, 247-263.

4. Harris, C. W., & Kaiser, H. F. Oblique factor analytic solutions by orthogonal transformations. Psychometrika, 1964, 29(4), 347-362.

5. Harris, M. L., & Harris, C. W. A factor analytic interpretation strategy. Educational and Psychological Measurement, 1971, 31, 589-606.

6. Harris, M. L., & Harris, C. W. Evidence concerning the Thurstone, Guttman, Guilford schema for classifying selected cognitive abilities. Paper presented at the meeting of the American Educational Research Association, New York City, February 1971.

7. Harris, M. L., & Harris, C. W. A structure of concept attainment abilities. Early draft of monograph, December 1972.

8. Kaiser, H. F., & Caffrey, J. Alpha factor analysis. Psychometrika, 1965, 30(1), 1-14.

9. Pennell, R. Routinely computable confidence intervals for factor loadings using the "jack-knife." The British Journal of Mathematical and Statistical Psychology, 1972, 25, 107-114.

10. Rao, C. R. Estimation and tests of significance in factor analysis. Psychometrika, 1955, 20, 93-111.

11. Tucker, L. R. An inter-battery method of factor analysis. *Psychometrika*, 1958, *23*, 111-136.

12. Tucker, L. R. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 1966, *31*, 279-311.

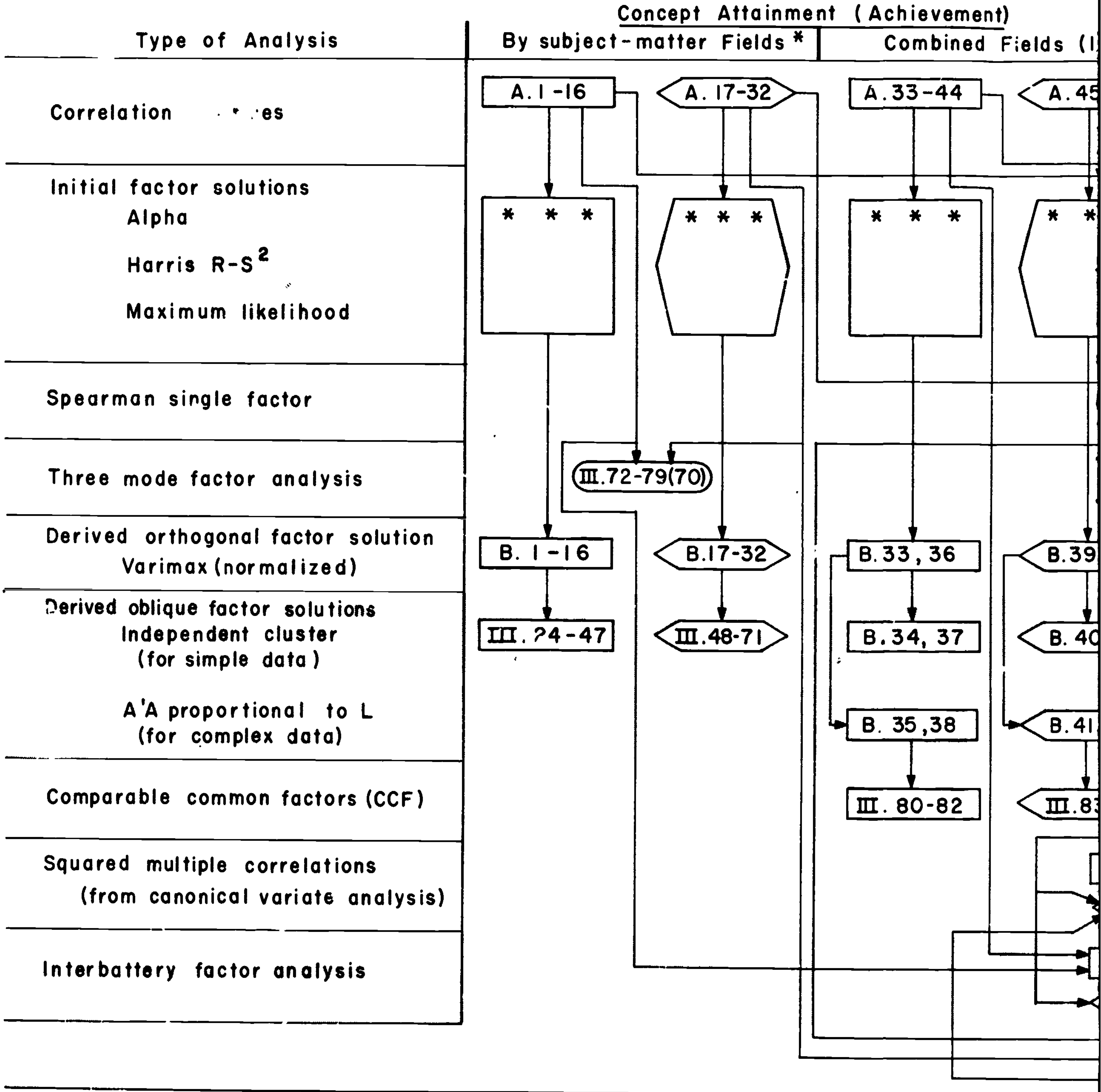
13. Tukey, J. W. Data analysis and behavioral scien[illegible] [illegible]ished manuscript, 1962.

14. Wingersky, B. G. Completeness, consistency and reliability in factor analysis. Unpublished manuscript, 1973.

AERA Symposium, February 26, 1973, New Orleans

GUIDE TO MULTIVARIATE METHODOLO

All entries in the boxes refer to table numbers (of the text and appendices of the draf



* For both 1970 and '971 unless a single year is indicated in parentheses.

** Selection of 31 tests for 1971 Study from 56 tests in 1970 Study.

*** Except for the number of factors, the initial factor solutions were not given in the draft monograph.

Harry H. Harman

LTIVARIATE METHODOLOGY OF THE PROJECT

ext and appendices of the draft monograph) where the results of the analysis will be found.